Rad s tabličnim podacima

Denis Vlašiček
Odsjek za psihologiju FFZG | CROSSDA
dvlasice@ffzg.hr
@dvlasicek
vdeni

Uvod

Uvod

Tech

Excel: Why using Microsoft's tool caused Covid-19 results to be lost

By Leo Kelion
Technology desk editor

5 October

MICROSOFT \ REPORT \ SCIENCE \

Scientists rename human genes to stop Microsoft Excel from misreading them as dates

Sometimes it's easier to rewrite genetics than update Excel

By James Vincent | Aug 6, 2020, 8:44am EDT

Kako se obraniti od ružnih podataka?

Kako se obraniti od ružnih podataka?

- što god radili, budite konzistentni
  - ▷ konzistentna imena varijabli
  - ▷ konzistentne oznake vrijednosti
  - ▷ konzistentna imena datoteka
  - ▷ konzistentan format datuma (`YYYY-MM-DD`; npr. `2020-12-04`)

Bevanda Denis - GLAVNI TAJNIK - predstavlja i zastupa: Pavliček Marijan - PREDSJEDNIK - predstavlja i zastupa
Žilavec Dubravko - PREDSJEDNIK - predstavlja i zastupa
Kosić Draženka - ČLAN UPRAVNOG TIJELA - predstavlja i zastupa: Cikota Zvjezdana - ČLAN STRANKE - predstavlja i
Lazar Zvjezdana - PREDSJEDNIK - predstavlja i zastupa
Bandić Milan - PREDSJEDNIK - predstavlja i zastupa: Kojić Slavko - POSLOVNI DIREKTOR - zastupa
Avdić Alija - PREDSJEDNIK - predstavlja i zastupa
Vesel Leonard - PREDSJEDNIK - predstavlja i zastupa
Josipović Ivo - PREDSJEDNIK - predstavlja i zastupa: Petek Renato - GLAVNI TAJNIK - zastupa
Raguž Željko - PREDSJEDNIK - predstavlja i zastupa
Karačić Branimir - PREDSJEDNIK IZVRŠNOG ODBORA - predstavlja i zastupa
Košiša Čičin-Šain Alenka - PREDSJEDNIK - predstavlja i zastupa
Puljak Marijana - PREDSJEDNIK - predstavlja i zastupa
Barun Mate - PREDSJEDNIK - predstavlja i zastupa
Balaž Kristijan - PREDSJEDNIK - predstavlja i zastupa
Harcet Mira - PREDSJEDNIK - predstavlja i zastupa: Kiš Arpad - GLAVNI TAJNIK - predstavlja i zastupa
Bobinac Nina - PREDSJEDNIK - predstavlja i zastupa
Kovar Martin - PREDSJEDNIK - predstavlja i zastupa

- u *jednu* ćeliju stavite samo *jedan* podatak

- ne ostavljajte prazne ćelije
  - ▷ ako vrijednost *treba* nedostajati, stavite eksplicitnu oznaku (na primjer `NA`)

NAZIV	SKRACENI_NAZIV	OSOBE	OIB	STATUS
MEDIMURSKA STRANKA	MS	Grašić Damir - PREDSJEDNIK - predstavlja i zastupa	36176331004	AKTIVNA
DEMOKRATSKA STRANKA UMIROVLJENIKA	DSU	Tičarić Milivoj - PREDSJEDNIK - predstavlja i zastupa	19578261879	AKTIVNA
ZELENA STRANKA	ZS	Babić Ališa - PREDSJEDNIK - predstavlja i zastupa	40857015687	AKTIVNA
JADRANSKA SOCIJALDEMOKRATSKA STRANKA HRVATSKE	JADRANSKI SOCIJALDEMOKRATI	Kovačić Boris - PREDSJEDNIK - predstavlja i zastupa; Marušić Josip - DOPREDSJEDNIK - predstavlja i zastupa	92762294306	AKTIVNA
DEMOKRATSKA KNEGINEČKA STRANKA	DKS	Keretić Damir - PREDSJEDNIK - predstavlja i zastupa	86378509070	AKTIVNA
AUTONOMNA REGIONALNA STRANKA HRVATSKOG PRIMORJA,	ARS	Jelušić Ivo - PREDSJEDNIK - predstavlja i zastupa	96925777632	AKTIVNA
AUTOHTONA - HRVATSKA STRANKA PRAVA	A - HSP	Keleminec Dražen - PREDSJEDNIK - predstavlja i zastupa	84404674327	AKTIVNA
ZELENI SAVEZ	ZELENI	Rupnik Josip Anton - PREDSJEDNIK - predstavlja i zastupa	58846771856	AKTIVNA
HRVATSKA STRANKA NEZAPOSLENIH	HSN	Milinović Viktor - PREDSJEDNIK - predstavlja i zastupa	82664540044	AKTIVNA
HRVATSKA RADNIČKA STRANKA	HRS	Novosel Mladen - PREDSJEDNIK - predstavlja i zastupa	36281215365	AKTIVNA
ZAGORSKA STRANKA	ZS	Jernečić Miljenko - PREDSJEDNIK - predstavlja i zastupa	80505604089	AKTIVNA
DEMOKRATSKA STRANKA ŽENA	DSŽ	Jelinčić Marija - PREDSJEDNIK - predstavlja i zastupa	27223591444	AKTIVNA
SOCIJALISTIČKA PARTIJA HRVATSKE	SPH	Šipov Dejan - PREDSJEDNIK - predstavlja i zastupa	18440010529	AKTIVNA
AKCIJA MLADIH	AM	Rubeša Vjekoslav - PREDSJEDNIK - predstavlja i zastupa; Buković Darjan - POTPREDSJEDNIK - predstavlja i zastupa;	17604594468	AKTIVNA
MEDIMURSKI DEMOKRATSKI SAVEZ	MDS	Pavlić Željko - PREDSJEDNIK - predstavlja i zastupa	22716317556	AKTIVNA
ALTERNATIVA		Simonić Bojan - PREDSJEDNIK - predstavlja i zastupa		AKTIVNA
HRVATSKI DEMOKRATSKI SAVEZ SLAVONIJE I BARANJE	HDSSB	Glavaš Branimir - PREDSJEDNIK - predstavlja i zastupa	37050651240	AKTIVNA
LISTA ZA RIJEKU	RI	Švorinić Danko - PREDSJEDNIK - predstavlja i zastupa; Basic Sandi - DOPREDSJEDNIK - predstavlja i zastupa	3376628315	AKTIVNA
HRVATSKA SELJAČKA STRANKA BRAĆE RADIĆ	HSS BRAĆE RADIĆ	Grčić Stanko - PREDSJEDNIK - predstavlja i zastupa	36114066915	AKTIVNA
HRVATSKA STRANKA PRAVNE DRŽAVE	HSPD	Prpić Siniša - PREDSJEDNIK - predstavlja i zastupa	56694372533	AKTIVNA
JEDINO HRVATSKA - POKRET ZA HRVATSKU	JEDINO HRVATSKA	Šibl Milovan - PREDSJEDNIK - predstavlja i zastupa	43647783176	AKTIVNA
SPLITSKA STRANKA UMIROVLJENIKA	SSU	Kovačić Ratko - PREDSJEDNIK - predstavlja i zastupa	18690517883	AKTIVNA
BOŠNJAČKA DEMOKRATSKA STRANKA HRVATSKE	BDSH	Hodžić Nedžad - PREDSJEDNIK - predstavlja i zastupa	39202199526	AKTIVNA
POPRAVI GRAD - NOVA GENERACIJA	PG - NG	Zoričić Ante - PREDSJEDNIK - predstavlja i zastupa	90654406562	AKTIVNA
NOVA SRPSKA STRANKA	NSS	Ladarević Svetislav - PREDSJEDNIK - predstavlja i zastupa	77799192134	AKTIVNA
DEMOKRATSKA PARTIJA SRBA	DPS	Kojadinović Mitar - PREDSJEDNIK - predstavlja i zastupa	40735782120	AKTIVNA
PAŠKA STRANKA - PAG	PSP	Marčić Željko - PREDSJEDNIK - predstavlja i zastupa	9837792562	AKTIVNA

- ne koristite formatiranje ćelija za pohranu podataka

OSOBE
Kosor Darinko - PREDSJEDNIK - predstavlja i zastupa
Dodig Goran - PREDSJEDNIK - predstavlja i zastupa
Plenković Andrej - PREDSJEDNIK - predstavlja i zastupa
Beljak Krešo - PREDSJEDNIK - predstavlja i zastupa
Miletić Boris - PREDSJEDNIK - predstavlja i zastupa
Vasilić Darijo - PREDSJEDNIK - predstavlja i zastupa
Starčević Karlo - PREDSJEDNIK - predstavlja i zastupa
Srebreniković Mirsad - PREDSJEDNIK - predstavlja i zastupa
Vrdoljak Ivan - PREDSJEDNIK - predstavlja i zastupa
Kojić Zoran - PREDSJEDNIK - predstavlja i zastupa
Brekalo Mirko - PREDSJEDNIK - predstavlja i zastupa; Dilberović Branko - DOPREDSJEDNIK - predstavlja i zastupa; Patrun Ivan - DOPREDSJEDNIK - predstavlja i zastupa
Tribuha Davor - PREDSJEDNIK - predstavlja i zastupa
Vukas Nikša - POSLOVNI DIREKTOR - zastupa; Bernardić Davor - PREDSJEDNIK - predstavlja i zastupa
Martan Ivan - PREDSJEDNIK - predstavlja i zastupa
Paraga Dobroslav - PREDSJEDNIK - predstavlja i zastupa
Vujanić Stjepan - PREDSJEDNIK - predstavlja i zastupa
Pupovac Milorad - PREDSJEDNIK - predstavlja i zastupa; Rkman Bogdan - PREDSJEDNIK IZVRŠNOG ODBORA - predstavlja i zastupa

OSOBE
Kosor Darinko predstavlja i zastupa
Dodig Goran predstavlja i zastupa
Plenković Andrej predstavlja i zastupa
Beljak Krešo predstavlja i zastupa
Miletić Boris predstavlja i zastupa
Vasilić Darijo predstavlja i zastupa
Starčević Karlo predstavlja i zastupa
Srebreniković Mirsad predstavlja i zastupa
Vrdoljak Ivan predstavlja i zastupa
Kojić Zoran predstavlja i zastupa
Brekalo Mirko predstavlja i zastupa; Dilberović Branko predstavlja i zastupa; Patrun Ivan predstavlja i zastupa
Tribuha Davor predstavlja i zastupa
Vukas Nikša zastupa; Bernardić Davor predstavlja i zastupa
Martan Ivan predstavlja i zastupa
Paraga Dobroslav predstavlja i zastupa
Vujanić Stjepan predstavlja i zastupa
Pupovac Milorad predstavlja i zastupa; Rkman Bogdan predstavlja i zastupa

predsjednik	predsjednik_fun	dopredsjednik	dopredsjednik_fun	glavni_tajnik	glavni_tajnik_fun
Kosor_Darinko	predstavlja_i_zastupa	nema	nije_primjenjivo	nema	nije_primjenjivo
Vasilić_Darijo	predstavlja_i_zastupa	nema	nije_primjenjivo	nema	nije_primjenjivo
Brekalo_Mirko	predstavlja_i_zastupa	Dilberović_Branko;Patrun_Ivan	predstavlja_i_zastupa;predstavlja_i_zastupa	nema	nije_primjenjivo
Kovačić_Boris	predstavlja_i_zastupa	Marušić_Josip	predstavlja_i_zastupa	nema	nije_primjenjivo
Mlinarić_Mato	predstavlja_i_zastupa	nema	nije_primjenjivo	Prezimenko-Novotny_Vladimir	predstavlja_i_zastupa

» There are no good ideas.

– Rory Allen Philip Ferreira

- varijablama dajte smislene nazive
  - ▷ izbjegavajte posebne znakove (!, ?, @ itd.)
  - ▷ izbjegavajte razmake (zamijenite ih s _)
  - ▷ izbjegavajte dijakritičke znakove (ili pazite na *encoding*)

HRVATSKA SOCIJALNO - LIBERALNA STRANKA	HSLS	Trg Nikole Ž̌ubi̇ć̌a Zrinskog 17/I, Zagreb
HRVATSKA DEMOKR̂ÂNSKA STRANKA	HDS	Palmoti̇ć̌eva 13, Zagreb
HRVATSKA DEMOKRATSKA ZAJEDNICA	HDZ	Trg Ž̌rtava fȧšizma 4/III, Zagreb
HRVATSKA SELJȦĆKA STRANKA	HSS	Zvonimirova 17, Zagreb
ISTARSKI DEMOKRATSKI SABOR	IDS	Splitska 3, Pula
PRIMORSKO GORANSKI SAVEZ	PGS	Ciottina 19, Rijeka
HRVATSKA STRANKA PRAVA	HSP	Primorska 5, Zagreb
STRANKA DEMOKRATSKE AKCIJE HRVATSKE	SDA HRVATSKE	Mandali̇ć̌ina 17, Zagreb
HRVATSKA NARODNA STRANKA - LIBERALNI DEMOKRATI	HNS	Ulica Kneza Mislava 8, Zagreb
SRPSKA NARODNA STRANKA	SNS	Mȧž̌urani̇ć̌ev trg 3, Zagreb
HRVATSKI OSLOBODILȦČKI POKRET	HOP	Ž̌enoina 27, Zagreb
HRVATSKA ČISTA STRANKA PRAVA	HČSP	Trg kralja Tomislava 2, Velika Gorica
SOCIJALDEMOKRATSKA PARTIJA HRVATSKE	SDP	Iblerov trg 9, Zagreb
HRVATSKA DEMOKRATSKA SELJȦČKA STRANKA	HDSS	Strosmayerova 9, Ž̌akovec
HRVATSKA STRANKA PRAVA - 1861	HSP - 1861	Remetinėć̌ka cesta 77a, Zagreb
ABECEDA DEMOKRACIJE	ABECEDA	Polj̇ć̌ka 13, Zagreb
SAMOSTALNA DEMOKRATSKA SRPSKA STRANKA	SDSS	Trg Drvena pijaca 28, Vukovar
HRVATSKA STRANKA UMIROVLJENIKA	HSU	Baruna Trenka 7, Zagreb
RAPSKI PU̇ČKI SABOR	RPS	Jurja Barakovi̇ć̌a 11, Rab
ZAGORSKA DEMOKRATSKA STRANKA	ZDS	Matije Gupca 53, Zabok
GRȦĐANSKA STRANKA SISAK	GSS	Ivana Mėđ̌trovi̇ć̌a 32, Sisak
SOCIJALISTİČKA RADNİČKA PARTIJA HRVATSKE	SRP	Pavla Hatza 14, Zagreb
DEMOKRATSKA PRIGORSKO-ZAGREBȦČKA STRANKA	DPS	Trg Ante Staṙč̌evi̇ć̌a 13, Sveti Ivan Zelina

- *nikakve* analize ili promjene u datotekama sa sirovim podacima
- napravite sigurnosne kopije
- napravite metapodatke
  - ▷ ime varijable
  - ▷ objašnjenje varijable
  - ▷ raspon vrijednosti

- da sve ne bude tužno

sex	age	years_service	n_children	work_satisf	elections_2020
f	48	36	1	54.24	HRVATSKE HRVATSKA ZAGREB LOZA
m	39	37	NA	63.28	VIŠNJAN STRANKA STRANKA PRAVA
f	NA	15	1	64.93	HRVATSKE HRVATSKA ZAGREB LOZA
m	43	34	NA	61.65	NA
f	24	9	1	NA	VIŠNJAN STRANKA STRANKA PRAVA
f	22	13	1	67.78	RAST PRAVA KAJKAVACA UNIJA
f	22	18	6	58.44	STRANKA NEZAVISNIH HRVATSKA
f	30	23	1	57.64	RAST PRAVA KAJKAVACA UNIJA
m	46	26	2	57.97	STRANKA NEZAVISNIH HRVATSKA
f	22	7	3	56.15	STRANKA NEZAVISNIH HRVATSKA
f	54	1	3	60.78	RAST PRAVA KAJKAVACA UNIJA
m	32	34	3	58.32	VIŠNJAN STRANKA STRANKA PRAVA
f	32	38	1	65.01	RAST PRAVA KAJKAVACA UNIJA
NA	27	30	NA	63.43	STRANKA NEZAVISNIH HRVATSKA
m	21	26	5	74.98	STRANKA NEZAVISNIH HRVATSKA
NA	50	14	1	65.07	VIŠNJAN STRANKA STRANKA PRAVA
m	21	18	1	66.39	VIŠNJAN STRANKA STRANKA PRAVA
f	50	27	5	NA	HRVATSKE HRVATSKA ZAGREB LOZA
f	34	25	1	63.21	STRANKA NEZAVISNIH HRVATSKA
f	22	10	1	66	STRANKA NEZAVISNIH HRVATSKA
m	33	30	2	66.41	HRVATSKE HRVATSKA ZAGREB LOZA
f	26	5	1	61.23	STRANKA NEZAVISNIH HRVATSKA
f	42	9	2	62.22	RAST PRAVA KAJKAVACA UNIJA
f	34	29	2	59.04	STRANKA NEZAVISNIH HRVATSKA
m	45	32	1	55.4	STRANKA NEZAVISNIH HRVATSKA
m	32	NA	2	61.51	HRVATSKE HRVATSKA ZAGREB LOZA
m	NA	8	1	69.49	STRANKA NEZAVISNIH HRVATSKA
f	31	33	1	61.94	RAST PRAVA KAJKAVACA UNIJA
m	44	7	2	62.72	STRANKA NEZAVISNIH HRVATSKA
m	26	26	4	60.83	RAST PRAVA KAJKAVACA UNIJA
f	40	NA	2	66.17	HRVATSKE HRVATSKA ZAGREB LOZA
f	50	2	4	57.84	RAST PRAVA KAJKAVACA UNIJA
f	22	24	1	60.85	HRVATSKE HRVATSKA ZAGREB LOZA
f	37	31	4	63.28	STRANKA NEZAVISNIH HRVATSKA
m	47	27	1	54.87	HRVATSKE HRVATSKA ZAGREB LOZA

?

?

- Horor priče o podacima.
- Prepreke na koje ste nailazili u vlastitom radu?

Dodatni materijali i reference

- morževi preuzeti iz Kastelein, R. A., Zweypfenning, R. C. V. J., Spekreijse, H., Dubbeldam, J. L., i Born, E. W. (1993). The Anatomy of the walrus head (Odobenus rosmarus). Part 3: The eyes and their function in Walrus ecology. *Aquatic mammals, 19*, 61-61.
- sadržaj temeljen na Broman, K. W., & Woo, K. H. (2018). Data organization in spreadsheets. *The American Statistician, 72*(1), 2-10. `https://doi.org/10.1080/00031305.2017.1375989`
- dodatna inspiracija iz CESSDA Training Team (2020). *CESSDA Data Management Expert Guide*. Bergen, Norway: CESSDA ERIC. doi: `https://doi.org/10.5281/zenodo.3820472`

Dodatni materijali i reference

- *Why using Microsoft's tool caused Covid-19 results to be lost*: `https://www.bbc.com/news/technology-54423988`
- *Scientists rename human genes to stop Microsoft Excel from misreading them as dates*: `https://www.theverge.com/2020/8/6/21355674/human-genes-rename-microsoft-excel-misreading-dates`